هاشمی بنجار، سیده الهام؛ ضیایی، ثریا؛ سلامی، مریم (۱۴۰۰). مدیریت کلان‌داده مبتنی بر رایانش‌ابری در کتابخانه‌های دانشگاه‌های سطح اول ایران.
پژوهش‌نامه کتابداری و اطلاع‌رسانی، ۱۱ (۱)، ۳۷۹- ۳۵۷. DOI: 10.22067/infosci.2021.67768.1001

مدیریت کلان‌داده مبتنی بر رایانش‌ابری در کتابخانه‌های دانشگاه‌های سطح اول ایران

سیده الهام هاشمی بنجار[۱]، ثریا ضیایی[۲]، مریم سلامی[۳]

تاریخ دریافت: ۹۹/۹/۲۴ تاریخ پذیرش: ۱۴۰۰/۱/۳۱ نوع مقاله: پژوهشی

چکیده

هدف: کلان‌داده مجموعه داده‌هایی است که با برخورداری از ویژگی‌هایی خاص، از طریق سیستم‌های نرم‌افزاری معمول امکان ذخیره‌سازی، مدیریت و پردازش آنها وجود ندارد. رایانش‌ابری مجموعه‌ای از منابع مجازی‌سازی شده با قابلیت دسترسی و استفاده آسان است. فناوری ذخیره‌سازی بر مبنای ابر قادر به مدیریت مؤثر کلان‌داده‌ها است. هدف این پژوهش تعیین وضعیت موجود مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش‌ابری و رتبه‌بندی مؤلفه‌های آن در کتابخانه‌های دانشگاه‌های سطح اول ایران است.

روش‌شناسی: پژوهش حاضر، از نوع مطالعات کاربردی و از نظر ماهیت، روش و کنترل متغیرها در زمره تحقیقات توصیفی است. جامعه آماری شامل کلیه کتابداران شاغل در کتابخانه‌های دانشگاه‌های سطح اول ایران به تعداد ۵۲۰ نفر است. پرسشنامه به‌صورت الکترونیکی در اختیار پاسخگویان قرار گرفت و در مجموع ۳۹۳ پرسشنامه دریافت و مورد تجزیه‌وتحلیل قرار گرفت. ابزار گردآوری داده‌ها پرسشنامه محقق ساخته که با بررسی مطالعات پیشین شش مولفه اصلی مؤثر بر مدیریت کلان داده‌های کتابخانه‌ای مبتنی بر رایانش‌ابری شامل نیروی انسانی، سازمان، زیر ساختار، مسائل اقتصادی، فرهنگ و مدیریت داده شناسایی شد که از طریق تحلیل عاملی تأییدی، روایی آن تأیید شد. برای تجزیه‌وتحلیل داده‌ها از آمار توصیفی و استنباطی و نرم‌افزارهای اس. پی. اس ۲۲[۴] و آموس ۲۲[۵] استفاده شد.

یافته‌ها: وضعیت موجود مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش‌ابری در کتابخانه‌های دانشگاه‌های سطح اول ایران از دیدگاه کتابداران، با میانگین ۲۳/۵۸۲ و انحراف استاندارد ۷/۶۵۵ معادل

۱. دانشجوی دکتری علم اطلاعات و دانش‌شناسی، دانشگاه پیام نور تهران، ایران، (نویسنده مسئول)، elnaz_hashemi2933@yahoo.com
۲. دانشیار گروه علم اطلاعات و دانش‌شناسی، دانشگاه پیام نور، تهران، ایران، Soraya.ziaei@gmail.com
۳. استادیار گروه علم اطلاعات و دانش‌شناسی، دانشگاه پیام نور، تهران، ایران، salamilib@yahoo.com

4. Spss
5. Amos

۰/۴۱۷۳۰–، نامطلوب ارزیابی شد. رتبه‌بندی مؤلفه‌های وضعیت موجود از قوت به ضعف به ترتیب ۱. نیروی انسانی با میزان بار عاملی ۶۴/.، ۲. مسائل اقتصادی با میزان بار عاملی ۰/۶۰، ۳. زیرساختار با میزان بار عاملی ۰/۵۹، ۴. مدیریت داده با میزان بار عاملی ۰/۵۸، ۵. فرهنگ با میزان بار عاملی ۰/۵۷، ۶. سازمان با میزان بار عاملی ۰/۵۲ قرار دارند.

نتیجه: کتابداران به واسطه درگیری روزمره‌شان با حجم بسیار بالای داده، عملا با کلان‌داده‌ها درگیر هستند. تطابق متوسط بودجه این کتابخانه‌ها با نیازهای فعلی و تأمین هزینه‌های لازم برای دیجیتال‌سازی منابع و وجود زیر ساختارهای قانونی سخت افزاری و نرم‌افزاری مناسب، راه را برای انجام اقدامات لازم برای بهره‌برداری از سرویس‌های ذخیره‌سازی مبتنی بر ابر را هموار می‌کند. اغلب کتابداران ضمن درک فرهنگ دسترسی آزاد به اطلاعات، اهمیت ارائه داده‌های تحلیل شده به مراجعین را درک کرده اما درک مراجعین از دریافت چنین داده‌هایی در حد متوسط می‌باشد. آگاهی و درک پایین مسئولین و عدم تمایل به تعامل و همکاری با سازمان‌های فعال در این حوزه، محیط پایداری را جهت کار با کلان‌داده‌ها و ارائه خدمات بر مبنای آنها را فراهم نمی‌کند.

کلیدواژه‌ها: مدیریت کلان داده، رایانش‌ابری، کتابخانه‌های دانشگاهی، کتابداران

مقدمه و بیان مسأله

در پی تغییرات در حوزه فناوری اطلاعات و ارتباطات و ظهور فناوری‌های نوین نظیر محاسبات ابری، اینترنت اشیاء، دستگاه‌های اسکن و شبکه‌های اجتماعی با حجم عظیمی از داده روبرو هستیم که در دنیای فناوری اطلاعات به‌عنوان کلان‌داده[1] شناخته می‌شوند. کلان‌داده‌ها به مجموعه داده‌هایی گفته می‌شود که توانایی دریافت، اکتساب، مدیریت و پردازش آنها در یک زمان قابل‌قبول به‌وسیله ابزارهای نرم‌افزاری و سخت‌افزاری سنتی وجود ندارد (چن، ماو، ژانگ و لیونگ[2]، ۱۳۹۴). از طرفی با توجه به رشد سریع کلان‌داده‌ها، ذخیره‌سازی، مدیریت و پردازش این حجم عظیم از داده‌ها نیازمند زیرساخت محاسباتی عظیم و پرهزینه و سخت‌افزارها و نرم‌افزارهای سنگین است. با مطرح بودن چالش‌هایی از قبیل لزوم پردازش سریع داده‌ها، هزینه‌های بالای راه‌اندازی، بررسی صحت داده‌ها، امنیت و سطوح دسترسی و بصری‌سازی، نیاز به مقیاس‌پذیری، تحمل خرابی، پاسخ به‌موقع، در دسترس بودن، کاهش هزینه‌ها و نیاز به ارتباطات میان تولیدکننده‌ها و مصرف‌کننده‌های داده در حوزه کلان‌داده، مناسب آن است که فراهم‌سازی این وظایف را به ابرها منتقل کرد. رایانش‌ابری[3] یک فناوری قدرتمند برای اجرای محاسبات سنگین و پیچیده است که نیاز به سخت‌افزارهای گران را حذف نموده و فضای محاسباتی و نرم‌افزار موردنیاز را در اختیار کاربر قرار

1. Big Data
2. Chen, Mao, Zhang & Leung
3. Cloud Computing

می‌دهد. آورام، لین، لین روآن و یه[1] فناوری رایانش‌ابری را الگوی جدیدی برای میزبانی و ارائه خدمات در اینترنت می‌دانند (Avram, 2014; Lin, Lin, Roan & Yeh, 2012). رایانش‌ابری علاوه بر حل چالش‌های مطرح شده برای کلان‌داده‌ها، دارای نکات مثبتی مانند تأمین منابع رایانشی از خارج از سازمان، قابل‌دستیابی بودن منابع رایانشی در اغلب موارد از طریق اینترنت، متغیر بودن هزینه استفاده از منابع بر اساس درخواست، مستقل بودن منابع رایانشی از وسیله، از دیدگاه کاربر نیز می‌باشد (ساسینسکی،۱۳۹۰). در حالی که می‌توان مزایای بی‌شماری را برای رایانش‌ابری نام برد، عیب‌های آن نیز کم نیستند. از معایب رایانش‌ابری می‌توان به مواردی از قبیل نیاز به اتصال دائمی به اینترنت، مشکلات امنیتی مانند سرقت اطلاعات و سرقت اکانت‌ها، کاهش کنترل سازمان بر داده‌ها اشاره کرد (همان منبع). در این میان کتابخانه‌ها به‌عنوان مخازن اطلاعاتی، سرمایه‌های دیجیتالی را ذخیره و مدیریت می‌کنند و عملکرد آنها از انبار اطلاعات به دسترسی به‌موقع به اطلاعات تغییر پیدا کرده است (Adebay, Ahmed & Adeniran, 2018; Qutab, Bhatti & Ullah, 2014).

در نگاه اول منابع چاپی و غیرچاپی کتابخانه‌ها در زمره کلان‌داده‌ها محسوب نمی‌شوند چراکه تمامی اطلاعات آماری، پژوهش‌ها، کتاب‌ها، مجلات با دیجیتال‌سازی به‌صورت ساختارمند قابل‌تولید توسط رایانه‌ها هستند؛ اما با نگاهی موشکافانه به منابع جدید اطلاعاتی بالفعل و بالقوه امروزی، داده‌های حاصل از منابع الکترونیکی پیوسته و ناپیوسته[2]، شبکه‌ای[3]، اینترنتی[4]، صوت و تصاویر، تراکنش‌های برخط، پست‌های الکترونیکی، داده‌های تلفن‌های همراه و برنامه‌های کاربردی آنها، دستگاه‌های تشخیص آر. اف.آی.دی[5]، شبکه‌های حسگر بی‌سیم، لاگ نرم‌افزارهای مختلف و تمامی داده‌های موجود در شبکه‌های اجتماعی می‌توانند منابع ارزشمندی برای کلان‌داده باشند. همچنین با در نظر گرفتن اطلاعات مربوط به انواع استنادها و ارجاعات به نویسندگان و مقالات و آثار مرتبط، فهرست منابع و تاریخ‌ها و بازنمایی مکانی با استفاده از سیستم اطلاعات جغرافیایی و بازنمایی مکانی مجموعه منابع کتابخانه در سطح قفسه، می‌توان با اطمینان اعلام کرد که چنین منابعی، کلان‌دادهایی تلقی می‌گردند که به‌منظور استخراج اطلاعات، کشف دانش و تصمیم‌گیری در خصوص مسائل مختلف کاربردی نیازمند مدیریت صحیح و برنامه‌ریزی‌شده هستند (Teets & Goldner,2013؛ چن و دیگران،۱۳۹۴). کلان‌داده‌ها در انواع متنی و غیرمتنی وجود دارند به همین دلیل

4. Avram, Lin, Lin, Roan & Yeh

۲. منابعی که روی سی. دی. رام یا فلاپی دیسک قابل‌دسترسی هستند و بر روی یک رایانه منفرد و مجزا مورداستفاده قرار می‌گیرند.

۳. منابع کتابشناختی یا منابع تمام متنی که در روی سرویس‌دهنده‌های شبکه‌ای ذخیره شده و از طریق شبکه سازمان قابل‌دسترسی هستند.

۴. منابع کتابشناختی یا منابع تمام متن که با استفاده از یکی از پروتکل‌های اینترنتی قابل‌دسترسی هستند (شهرزادی،۱۳۹۱).

5. RFID: Radio Frequency Identification

در مدیریت کلان‌داده‌های کتابخانه‌ای در مرحله اول مشابه با چهار وظیفه اصلی کتابخانه‌ها با جمع‌آوری، ذخیره‌سازی، جستجو، به اشتراک‌گذاری منابع روبرو بوده و سپس با تحلیل کلان‌داده‌ها، به استخراج اطلاعات ارزشمند و پیش‌بینی و تولید دانش نو و کاربردی، تعامل اطلاعات مفید، معنا بخشی (ایجاد بافت) و همچنین به تلاش برای رسیدن به آرمان دانش‌محوری در مسائل اقتصادی، سیاسی، اجتماعی و فرهنگی پرداخته می‌شود. همچنین به موازات تحولات فناوری‌های نوین در محیط اطلاعاتی و دسترسی از راه دور به شبکه‌ها و پایگاه‌های اطلاعاتی، مراجعه حضوری به کتابخانه‌ها کاهش یافته است. بسیاری از کاربران تمایلی به مراجعه حضوری به کتابخانه‌ها نداشته و مطالعه کتاب‌های الکترونیکی، جستجوی اینترنت و پایگاه‌های اطلاعاتی و خدمات غیرحضوری و فضاهای مجازی را بیشتر ترجیح می‌دهند. یک دیدگاه رایج در این باره وجود دارد که می‌گوید «اینترنت برای کسب موقعیت مدیریت دانش با کتابخانه‌های سنتی در حال رقابت است» (Li, Lu, Dou & Wang, 2017). در حال حاضر، بسیاری از کاربران معتقدند که کتابخانه‌های سنتی با اینترنت جایگزین شده و این مراکز موقعیت و جذابیت خود را از دست داده‌اند (Head, Komori, LaMere, Whisenant, Nieuwerburgh, Salomon, & Ordoukhanian,2014; Marshall, & Sollenberger, Easterby-Gannett, Morgan, Klem, Cavanaugh Hunter,2013).

کتابخانه‌ها با انتقال منابع و خدمات خود به محیط‌های مجازی و استفاده از فناوری‌های کلان‌داده و رایانش‌ابری قادر به تمرکز بر روی مأموریت اصلی خود یعنی ارائه خدمات اطلاع‌رسانی، همگام با نیازهای به‌روز کاربران خود خواهند بود. نقش کتابخانه‌های دانشگاهی به‌عنوان بخشی از نظام آموزش عالی که مرکز گردآوری و انتقال اطلاعات علمی و کانون تأمین امکانات تحقیق برای پژوهشگران، اساتید، دانشجویان و توسعه علوم می‌باشد، محرز است. امروزه کتابخانه‌های دانشگاهی در حال توسعه پایگاه‌ها و مخازن داده‌ای خود و کمک به پژوهشگران در نگارش طرح‌های مدیریت داده، مستندسازی و آموزش چگونگی استفاده از کلان‌داده‌ها هستند. همچنین کتابخانه‌هایی که منابع مالی و انسانی کافی ندارند به دنبال کشف راه‌هایی برای شروع این اقدام مهم هستند (Fister, 2015). محققان برای مدیریت داده‌ها چهار نقش مهم را ذکر کرده‌اند: داده‌ساز[1]، دانشمند داده، مدیر داده و کتابدار داده (Kim & choi, 2016, Lyon, Wright, Corti, Edmunds & Bennett, 2013, Pryor & Donnelly, 2009; Swan & Brown, 2008). مهارت‌هایی همچون مهارت مذاکره و برقراری ارتباط بین سایر نقش‌ها، بررسی و مدیریت و کنترل، ارتقاء و حمایت از بازاریابی سازمان، ارتقای آگاهی سازمانی و اشتراک‌گذاری دانش و اطلاعات، ارزیابی

1. Data Creator

داده و نگهداری از آن از نقش‌های منحصربه‌فرد کتابدار داده است که فقط خود این فرد می‌تواند از عهده آن برآید. ارزش‌گذاری داده‌ها در موضوع‌های اقتصادی و حفاظت از داده‌ها دو نقش مشترک کتابدار داده یا مدیر داده است (درمنده، نوروزی، اسمعیلی گیوی،۱۳۹۷). از مزایای بیان شده برای کلان‌داده در کتابخانه‌های دانشگاهی می‌توان به طرح‌های سازماندهی کلان‌داده، (Reinhalter & Wittmann,2014; Gordon-Murnane,2012;Witt,2012; Heidorn,2011)، شناخت نظری و عملی کلان‌داده و تجزیه‌وتحلیل آن در کتابخانه‌های دانشگاهی (Rani,2016)، حمایت از محققین (Keil, 2014)، سرمایه گذاری بر روی فرصت‌های ایجاد شده توسط کلان‌داده و روش‌های استخراج متون (Al-Daihani & Abrahams,2016) اشاره کرد.

اما ایفای این نقش به میزان کارآیی خدمات این کتابخانه‌ها بستگی دارد. برخورداری کتابخانه‌های دانشگاهی از به‌روزترین روندها و تجهیزات روزآمد با هدف ارائه خدمات و اشاعه اطلاعات امری مهم است. لذا ضرورت مدیریت کلان‌داده‌ها با در نظر گرفتن وظیفه اصلی و هدف کارکردی کتابخانه‌ها، بر عهده کتابخانه‌ها و کتابداران و متخصصین علم اطلاعات است. بدین ترتیب کتابداران می‌توانند علاوه بر انجام نقش‌های اصیل خود در جمع‌آوری، ذخیره‌سازی، سازماندهی و اشاعه صحیح داده‌ها؛ سکان‌دار مدیریت داده‌ها شوند. در این پژوهش ضمن مروری بر مفاهیم نظری این حوزه، با ارزیابی وضعیت موجود مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش‌ابری در کتابخانه‌های دانشگاه‌های سطح اول کشور، به شناخت شرایط فعلی و رتبه‌بندی مؤلفه‌های اثرگذار در مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش‌ابری پرداخته تا قدم در راه شروع بهره‌گیری از مزایای آن و همچنین شروع انقلاب کلان‌داده‌ای در کتابخانه‌ها برداشته شود. انتظار می‌رود این پژوهش برای آشنایی بیشتر پژوهشگران در حوزه علم اطلاعات و دانش‌شناسی با مبانی نظری و اجرای پژوهش‌های جدید و بررسی سایر زوایای این حوزه مفید باشد. همچنین می‌تواند برای سیاستگذاران و مدیران دانشگاه‌ها و کتابخانه‌ها و کتابداران که به‌دنبال برنامه‌ریزی، سیاستگذاری و بهره‌گیری از فناوری‌های کلان‌داده و مهاجرت به ابرها هستند، مفید فایده واقع شود.

کلان‌داده

برای کلان‌داده تعریف‌های متعددی ارائه شده است. «موسسه مک‌کینزی[1]» کلان‌داده را مرز بعدی برای نوآوری رقابت و بهره‌وری می‌داند و معتقد است که کلان‌داده‌ها باید مجموعه‌هایی را پوشش دهند که امکان گردآوری، ذخیره و مدیریت آنها به‌وسیله نرم‌افزارهای پایگاه داده کلاسیک وجود ندارد (Manyika,

1. McKinsey Global Institute

(Chui, Brown, Bughin, Dobbs, Roxburgh & Byers, 2011). موسسه ملی فناوری و استانداردها، کلان‌داده را شامل مجموعه‌های داده‌ای گسترده‌ای معرفی می‌کند که در درجه اول از نظر حجم[1]، سرعت[2] و تنوع[3] نیازمند معماری‌های مقیاس‌پذیر برای ذخیره‌سازی کارآمد، و تجزیه‌وتحلیل آنها است. کلان‌داده‌ها، حجم بالایی از داده‌های غیرساختاریافته[4] هستند که با سیستم‌های مدیریت داده‌های استاندارد، قابل مدیریت و بهره‌برداری نیستند (Gantz & Reinsel, 2011; Kaisler, Armour, Espinosa, & Money, 2013). داده‌های موجود در دنیای امروز را می‌توان به ۳ بخش داده‌های ساختاریافته[5]، نیمه ساختاریافته[6] و ساختارنیافته تقسیم کرد (ارل، ختاک و بوهلر[7]، ۱۳۹۵). داده‌های ساختاریافته، داده‌هایی هستند که دارای قالبی منظم و قابل پیش‌بینی بوده و به‌وسیله یک زیرساخت مناسب مدیریت می‌شوند. این نوع داده‌ها قابل درک برای موتورهای جستجو هستند. داده‌های نیمه ساختاریافته، داده‌هایی هستند که ساختار رسمی «مدل داده» را ندارند، یعنی فاقد تعریف جدول در یک پایگاه داده رابطه‌ای هستند. فرمت‌های داده ایکس. ام. ال.[8] و مستندات جی سون[9] دو نوع از متداول‌ترین داده‌های نیمه‌ساختاریافته هستند. داده‌های بدون ساختار، داده‌هایی هستند که شکل و ساختار مشخصی ندارند. داده‌های بدون ساختار نیز غیرقابل‌پیش‌بینی بوده و سرعت رشد این نوع از داده‌ها در مقایسه با دو نوع دیگر، بیشتر است. از متداول‌ترین آنها می‌توان به فایل‌های متنی، محتواهای چندرسانه‌ای مانند تصاویر و فایل‌های صوتی اشاره کرد. سه ویژگی حجم، سرعت و تنوع به‌عنوان ویژگی‌های اصلی کلان‌داده معرفی شده است (Laney,2001). اما بعدها و با پیشرفت‌های متعدد در این حوزه ویژگی‌هایی از قبیل صحت[10]، ارزش[11]، نوسان[12]، اعتبار[13] و نمایش داده‌ها[14] مطرح گردید. از چالش‌های مطرح شده برای کلان‌داده‌ها می‌توان به فشردگی و کاهش افزونگی داده‌ها[15]، کشف دانش و

1. Volume
2. Velocity
3. Variety
4. Unstructured
6. Structured data
7. Semi Structured data
7. Erl, Khattak & Buhler
8. XML: Extensible Markup Language
9. JSON
10. Veracity
11. Value
12. Volatility
13. Validity
14. Data Representation
15. Redundancy Reduction and Data Compression

پیچیدگی‌های محاسباتی، محرمانگی داده‌ها[1]، مدیریت انرژی، توسعه‌پذیری و مقیاس‌پذیری[2] و همکاری اشاره کرد (چن و دیگران،۱۳۹۴).

رایانش‌ابری

موسسه ملی فناوری و استاندارد امریکا رایانش‌ابری را این‌گونه تعریف می‌کند: رایانش‌ابری مدلی است برای ایجاد دسترسی همگانی و به‌صورت راحت و آسان و بر اساس سفارش، به مجموعه‌ای از منابع رایانشی پیکربندی‌پذیر مانند شبکه‌ها، سرویس‌دهنده‌ها، فضای ذخیره‌سازی، برنامه‌های کاربردی و سرویس‌هایی که بتوانند با کمترین نیاز به تلاش مدیریتی یا تعامل با ارائه‌دهنده سرویس، به سرعت تخصیص‌یافته یا آزاد شوند. این مدل ابری پنج مشخصه اساسی شامل ۱. خدمات مستقیم به محض نیاز[3]، ۲. دسترسی به شبکه با پهنای باند وسیع[4]، ۳. منابع تجمیعی[5]، ۴. انعطاف‌پذیری سریع[6]، ۵. خدمات یا سرویس اندازه‌گیری شده[7] دارد و از سه مدل سرویس‌دهی شامل نرم‌افزار به‌عنوان یک سرویس[8]، بستر ابری به‌عنوان سرویس[9]، زیرساخت ابری به‌عنوان سرویس[10] و چهار مدل پیاده‌سازی شامل ابر خصوصی[11]، ابر عمومی[12]، ابر گروهی[13]، ابر آمیخته[14] تشکیل شده است.

پیشینه پژوهش

حمد، فخوری و عبدل جبار[15] (۲۰۲۰)، در پژوهشی با عنوان «فرصت‌ها و چالش‌های کلان‌داده برای تجزیه‌وتحلیل راهبردها در کتابخانه‌های دانشگاهی اردن» نشان دادند که کارکنان واحد فناوری اطلاعات در سه کتابخانه اصلی دانشگاه اردن با مفهوم کلان‌داده کاملاً آشنا هستند و این آگاهی به کتابخانه‌های دانشگاهی منجر به اخذ تصمیمات نوآورانه و مقرون‌به‌صرفه می‌گردد.

1. Data Confidentiality
2. Expendability and Scalability
3. On-Demand Self-Service
4. Broad Network Access
5. Resource Pooling
6. Rapid Elasticity
7. Measured Service
8. SaaS: Cloud Software as a Service
9. PaaS: Cloud Platform as a Service
10. IaaS: Cloud Infrastructure as a Service
11. Privet Cloud
12. Public Cloud
13. Community Cloud
14. Hybrid Cloud
15. Hamad, Fakhuri & Abdel Jabbar

بال[1] (۲۰۱۹)، در پژوهشی با عنوان «کلان‌داده و تأثیرات آن بر کتابخانه‌ها» به مرور امکانات و فرصت‌های استفاده از حجم بسیار زیاد داده‌ها در کتابخانه با بیان فرضیات و مثال‌هایی کاربردی پرداخته است. نویسنده برای به‌کارگیری کلان‌داده‌ها در کتابخانه‌های جهان مثال‌هایی را مطرح می‌کند که از میان آنها می‌توان به کتابخانه کنگره: ایجاد و تهیه بایگانی توییتر، ایجاد پایگاه فراداده برای داده‌های ژئوفیزیک در استرالیا، برنامه کاربردی کلان‌داده برای کتاب‌ها (کتابخانه دانشگاه‌هاروارد)، پروژه آزمایشگاه‌های کتابخانه‌ای کمیته مشترک سیستم‌های اطلاعاتی[2] و آژانس آمار آموزش عالی بریتانیا[3] و موارد دیگر اشاره کرد.

یوهانگ[4] (۲۰۱۸)، در پژوهشی با عنوان «بهبود کارایی و بازسازی منابع دیجیتال در کتابخانه‌های دانشگاه از منظر کلان‌داده‌ها بر اساس آمار استفاده از منابع دیجیتال در کتابخانه اتحادیه دانشگاه چانگبی» با اشاره به سرعت بالای پردازش کلان‌داده‌ها و جنبه‌های با ارزش آن، وجود کلان‌داده را به‌مثابه کانالی برای استفاده از منابع دیجیتالی می‌داند. بنابراین کاربران قادر به صرفه‌جویی در مدت زمان جستجو و بازیابی منابع هستند.

درمنده، نوروزی و اسمعیلی گیوی (۱۳۹۷) در مقاله‌ای با عنوان «فرصت‌های مدیریت داده‌های بزرگ در کتابخانه‌ها و مراکز اطلاع‌رسانی: واکاوی ساختاری-تفسیری و ارائه راهکار» با هدف شناسایی، تعیین، اولویت‌بندی و تحلیل فرصت‌های مدیریت داده‌های بزرگ و واکاوی ساختاری-تفسیری آن به ارائه راهکارهایی برای مدیریت کلان‌داده‌ها در کتابخانه‌ها و مراکز اطلاع‌رسانی شهر تهران پرداخته‌اند. یافته‌های این پژوهش نشان داد که فرصت‌های مدیریت کلان‌داده از نظر مدیران مراکز اطلاع‌رسانی شهر تهران به ترتیب اهمیت و اولویت عبارتند از: مدیریتی-سازمانی، ماهیتی، فرآیندی، انسانی و از دید خبرگان عبارتند از: ماهیتی، فرایندی، انسانی و مدیریتی-سازمانی.

آقایی میرک آباد، آهنگر، آقایی میرک آباد و حسن‌پور (۱۳۹۵) در مقاله‌ای با عنوان «سیستم‌های مدیریت کلان‌داده‌ای مبتنی بر مدل ابری» با هدف ارائه مدلی پیشنهادی برای سیستم‌های مدیریت کتابخانه کلان‌داده‌ای با روش کتابخانه‌ای و تحلیل متون، ضمن مروری بر مفهوم و قابلیت‌های کلان‌داده‌ها و نیز شرح مختصری از ارزش ذخیره و بازیابی کلان‌داده‌ها به توضیح روش داده‌کاوی در حد کلان بر اساس ماتریس تصادفی برای دسترسی سریع‌تر و بهتر به اطلاعات، به‌عنوان یک ایده پرداخته و این نوع مدل کتابخانه ابری با کلان‌داده را به‌عنوان راه‌حلی برای برخی مشکلات کتابخانه‌های دیجیتال مانند دسترسی سریع به اطلاعات،

1. Ball
2. JISC: Joint Information System Committee
3. HESA: the British Higher Education Statistics Agency
4.Yuhong

پیشنهاد می‌کند.

صراف‌زاده (۱۳۹۴) در پژوهشی با عنوان «مدیریت کلان‌داده‌های پژوهشی: نقشی نوین برای کتابخانه‌های دانشگاهی» با نگاهی به ابعاد مختلف مدیریت داده‌های پژوهشی و نقش‌های بالقوه‌ای که کتابخانه‌ها و کتابداران دانشگاهی می‌توانند در این زمینه‌ایفا کنند، بیان می‌دارد که کمک به مدیریت داده‌های پژوهشی را نقشی جدید و چالشی برای کتابخانه‌ها است که دانش و مهارت‌های جدید را می‌طلبد. از برخی مصداق‌های این مهارت‌ها برای مدیریت داده‌های پژوهشی می‌توان به آمار، آشنایی با ابزارها و فنون کتاب‌سنجی و علم‌سنجی، تحلیل استنادی، ضریب تأثیرها و... برای ارزیابی پژوهش‌ها، دانش موضوعی، مهارت حفاظت از داده‌ها، مهارت‌های فناوری اطلاعات و ارتباطات و... اشاره کرد.

جمع‌بندی پیشینه پژوهش

مطالعه و بررسی پژوهش‌های انجام شده نشان داد که تاکنون پژوهش مشخصی که به بررسی وضعیت مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌های دانشگاهی ایران بپردازد، انجام نشده است. پژوهش‌های خارج از کشور به مباحثی از قبیل مدیریت کلان‌داده‌ها در کتابخانه‌ها، مدیریت داده و حفظ حریم خصوصی، تعاریف و تاریخچه کلان‌داده‌ها، تجربه‌های کتابداران، فرصت‌ها و چالش‌های پیش روی کتابداران در کار با کلان‌داده‌ها، عناوین جدید شغلی و برنامه‌های آموزشی در کار با کلان‌داده‌ها پرداخته‌اند. در پژوهش‌های داخلی انجام شده نیز به مباحثی از قبیل ارائه تعریف‌ها و فرصت‌های به‌کارگیری کلان‌داده‌ها در کتابخانه‌ها و ارائه راهکارهایی برای مدیریت کلان‌داده‌ها در کتابخانه‌ها، مدیریت کلان‌داده‌های پژوهشی و نقش کتابداران در مدیریت و یافت‌پذیری کلان‌داده‌ها پرداخته شده است. بنابراین پژوهشی که به مدیریت کلان‌داده‌های کتابخانه‌ای به‌همراه به‌کارگیری راه‌حل‌های رایانش ابری و بررسی چگونگی وضعیت این موضوع در کتابخانه‌ها بپردازد، انجام نشده است.

پرسش‌ها

پرسش اول: آیا مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌های دانشگاه‌های سطح اول ایران از دیدگاه کتابداران از وضعیت مطلوبی برخوردار است؟

پرسش دوم: رتبه‌بندی مؤلفه‌های مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌های دانشگاه‌های سطح اول ایران از دیدگاه کتابداران برای وضعیت موجود چگونه است؟

روش‌شناسی

پژوهش حاضر، از نوع مطالعات کاربردی و از نظر ماهیت، روش و کنترل متغیرها در زمره تحقیقات

توصیفی می‌باشد. جامعه آماری پژوهش حاضر شامل کلیه کتابداران شاغل در کتابخانه‌های دانشگاه‌های سطح اول ایران به تعداد ۵۲۰ نفر می‌باشد. دانشگاه‌های سطح اول ایران بر اساس اجرای ماده ۳ سیاست‌های و ضوابط اجرایی حاکم بر آمایش آموزش عالی موسوم به طرح آمایش مصوب ۹۴/۱۲/۱۸ شورای عالی انقلاب فرهنگی، حسب تصویب ستاد آمایش آموزش عالی در دو دسته جامع و تخصصی، شامل ۱۳ دانشگاه، در سایت وزارت علوم تحقیقات و فناوری به تاریخ ۹۶/۷/۸ و به شماره ۲/۱۷۸۹۷۸ می‌باشد[1]. با توجه به محدود بودن پاسخ‌دهندگان و در جهت تلاش برای کسب داده‌های واقعی، نمونه‌گیری انجام نگرفت و همه کتابداران شاغل در کتابخانه‌های دانشگاه‌های سطح اول ایران لحاظ شدند. پرسشنامه به‌صورت الکترونیکی و به‌صورت یکسان برای همه افراد ارسال شد. جهت دستیابی به اهداف تحقیق برای شناسایی وضعیت موجود از پرسشنامه محقق ساخته استفاده شد. برای تهیه آن ابتدا با مطالعات بسیار در منابع فارسی و خارجی و مطالعه نظرات خبرگان در حوزه فناوری اطلاعات و علم اطلاعات و دانش‌شناسی لیستی از موارد مورد نظر تهیه و بعد از مشورت تنی چند از اساتید و صاحب‌نظران در دو حوزه فناوری اطلاعات و علم اطلاعات و دانش‌شناسی با استفاده از فن دلفی اقدام به یکدست‌سازی و دسته‌بندی آنها در گروه‌های عام‌تر شد. سپس برای اطمینان از این موضوع که سؤالات مطرح شده دقیقاً به بررسی وضعیت موجود می‌پردازد، تحلیل عاملی تأییدی انجام شد.

تحلیل عاملی تأییدی مرتبه اول

قبل از انجام تحلیل عاملی جهت اطمینان از کفایت تعداد داده‌های موجود از شاخص کا.ام. او.[2] استفاده شد (KMO:۰/۹۷۴). بنابراین حجم نمونه برای انجام تحلیل عاملی کافی است. برای داشتن روایی باید بین مؤلفه‌ها و سؤالات پرسشنامه، همبستگی معناداری وجود داشته باشد. بار عاملی بین ۰/۳ تا ۰/۶ قابل‌قبول است (حبیبی، ۱۳۹۶). مقادیر تی، معنادار بودن هر یک از مؤلفه‌ها را نشان می‌دهد و چنانچه مقدار تی، بزرگتر از قدر مطلق عدد ۱/۹۶ باشد، مؤلفه‌های مدل معنادار هستند. تمامی بارهای عاملی برای سؤالات مربوط به هر مؤلفه در پرسشنامه، بیشتر از ۰/۴ و مقادیر تی بیشتر از ۱/۹۶ به‌دست آمد. لذا اعتبار پرسشنامه تأیید شد. آلفای کرونباخ تمامی مؤلفه‌های وضعیت موجود بیشتر از ۰/۷ برآورد شد. (آلفای کرونباخ برای مؤلفه نیروی انسانی برابر با ۰/۹۸۶، آلفای کرونباخ برای مؤلفه سازمان برابر با ۰/۹۶۷، آلفای کرونباخ برای مؤلفه زیرساختار برابر با ۰/۹۵۲، آلفای کرونباخ برای مؤلفه مسائل اقتصادی برابر با ۰/۹۳۳، آلفای کرونباخ

1. https://edu.msrt.ir/fa/news/35266/ابلاغ-نتایج-سطح-بندی-دانشگاه-های-دولتی
2. Kaiser-Mayer-Olkin

برای مؤلفه فرهنگ برابر با ۰/۹۵۷ و آلفای کرونباخ برای مؤلفه مدیریت داده برابر با ۰/۹۷۶)؛ لذا نیاز به حذف هیچ مؤلفه‌ای از مدل نبود. به‌طورکلی، برای ارزیابی مدل تحلیل عاملی تأییدی چندین مشخصه برازندگی وجود دارد. در این پژوهش برای ارزیابی مدل تحلیل عاملی تأییدی از شاخص‌های کای اسکوئر (x2)، شاخص نیکویی برازش (GFI)، شاخص برازش مقایسه‌ای (CFI)، شاخص برازش فزاینده (IFI) و شاخص بسیار مهم ریشه دوم برآورد واریانس خطای تقریب یا همان ریشه میانگین مربعات خطای برآورد (RMSEA) استفاده شد. در پژوهش حاضر مقدار کای دو به درجه آزادی ۱/۱۳ و کمتر از ۳ بود. مقدار جذر برآورد واریانس خطای تقریب (RMSEA) برابر با ۰/۰۱۸ و کمتر از ۰/۰۸ بود. شاخص برازندگی تطبیقی (CFI) برابر با ۰/۹۸۷، شاخص برازندگی افزایشی (IFI) برابر با ۰/۹۸۷ و شاخص نیکویی برازش (GFI) برابر با ۰/۸۲۸ محاسبه شد. بنابراین در حالت کلی و با توجه به شاخص‌های محاسبه شده می‌توان برازش مطلوب مدل را نتیجه گرفت و از آنها برای پاسخ به پرسش‌های پژوهش استفاده کرد. در نهایت پرسشنامه شناسایی وضعیت موجود با ۶ مؤلفه، شامل مؤلفه نیروی انسانی (۲۵ سؤال)، سازمان (۱۴ سؤال)، زیر ساختار (۸ سؤال)، مسائل اقتصادی (۶ سؤال)، فرهنگ (۱۱ سؤال) و مدیریت داده (۱۷ سؤال) به شناسایی وضعیت موجود مدیریت کلان داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌های دانشگاه‌های سطح اول ایران با استفاده از طیف پنج‌تایی لیکرت (خیلی زیاد، زیاد، نمی‌دانم، کم و خیلی کم) آماده شد. پرسشنامه به‌صورت الکترونیکی در میان جامعه آماری که تعداد کل آنها بر اساس بررسی‌های محقق و شناسایی کل جامعه پژوهش تعداد ۵۲۰ نفر بود؛ توزیع شد که در مجموع ۳۹۳ پرسشنامه دریافت و مورد تجزیه‌وتحلیل قرار گرفت. در ادامه و بعد از شناسایی وضعیت موجود، با استفاده از تحلیل عاملی تأییدی مرتبه دوم، مؤلفه‌های مؤثر در شناخت وضعیت موجود رتبه‌بندی و نتایج پس از تجزیه‌وتحلیل، به‌طور توصیفی ارائه شد.

یافته‌ها

ویژگی‌های فردی و حرفه‌ای پاسخگویان

بر اساس یافته‌های مرتبط با ویژگی‌های فردی و حرفه‌ای پاسخگویان، بیشترین تعداد ۲۴۳ نفر یا ۶۱/۸ درصد پاسخگویان زن می‌باشند. از نظر توزیع سن، بیشترین تعداد ۴۷ نفر یا ۱۲ درصد پاسخگویان دارای سن ۳۹ سال و کمترین تعداد معادل یک نفر برای سنین ۲۵ و ۴۷ سال می‌باشند. بیشترین تعداد ۱۹۳ نفر یا ۴۹/۱ درصد پاسخگویان مدرک تحصیلی کارشناسی ارشد و کمترین تعداد ۳۹ نفر مدرک تحصیلی کاردانی یا پایین‌تر از آن‌را دارند. بیشترین تعداد ۲۶۲ نفر معادل ۶۶/۷ درصد در رشته علم اطلاعات و

دانش‌شناسی و کمترین تعداد یعنی ۲۶ نفر معادل ۶/۶ درصد پاسخگویان در رشته‌های فناوری اطلاعات و ۱۰۵ نفر معادل ۲۶/۷ درصد در سایر رشته‌ها تحصیل کرده‌اند. بیشترین تعداد پاسخگویان از دانشگاه تهران،۶۲ نفر معادل ۱۵/۸ درصد و ۵۴ نفر از دانشگاه فردوسی مشهد معادل ۱۳/۷ درصد و کمترین تعداد ۱۴ نفر معادل ۳/۶ درصد از دانشگاه صنعتی خواجه‌نصیرالدین طوسی بوده‌اند.

پاسخ به پرسش اول: آیا مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌های دانشگاه‌های سطح اول ایران از دیدگاه کتابداران از وضعیت مطلوبی برخوردار است؟

آزمون تی تک نمونه‌ای[1]

با توجه به تأیید نرمال بودن توزیع داده‌ها از طریق اجرای آزمون کولموگروف‌اسمیرنوف[2] (۰/۰۵۵ =p و z=۰/۰۴۵) و وجود مقدار مفروض و نظری ثابت یعنی میانگین به‌دست آمده از میانگین مؤلفه وضعیت موجود یعنی عدد ۲۴/۵ آزمون تی تک نمونه‌ای برای تعیین ارزیابی وضعیت موجود انجام شد.

جدول ۱. آزمون تی تک نمونه‌ای برای مقایسه نمره وضعیت موجود با نمره معیار

متغیر	میانگین	انحراف استاندارد	نمره معیار	t	p	اختلاف میانگین
وضعیت موجود	۲۳/۵۸۲	۷/۶۵۵	۲۴/۵	۲/۳۷۵-	۰/۰۱۸	۰/۹۱۷-

مطابق نتایج حاصل از جدول ۱، میانگین نمره وضعیت موجود در گروه نمونه ۲۳/۵۸۲ با انحراف استاندارد ۷/۶۵۵ که به‌طور معناداری (۰/۰۵ > ۰/۰۱۸) معادل ۰/۹۱۷- از نمره معیار پایین‌تر بوده است. بنابراین میانگین نمره مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌های دانشگاه‌های سطح اول کشور به‌طور معنی‌داری در وضعیت نامطلوب قرار دارند.

پاسخ به پرسش دوم: رتبه‌بندی مؤلفه‌های مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌های دانشگاه‌های سطح اول ایران از دیدگاه کتابداران برای وضعیت موجود چگونه است؟

تحلیل عاملی تأییدی مرتبه دوم[3]

تحلیل عاملی تأییدی مرتبه دوم با هدف رتبه‌بندی مؤلفه‌های مؤثر در شناخت وضعیت موجود در مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش‌ابری در کتابخانه‌های دانشگاه‌های سطح اول انجام شد.

1. One Sample T-Test
2. Kolmogorov–Smirnov test
3. Second order confirmatory factor analysis

ضرایب استاندارد شده بارهای عاملی و مقدار معناداری تی

جدول ۲ ضرایب معناداری تی و ضرایب مسیر استاندارد شده بین وضعیت موجود و مؤلفه‌های آن‌را نشان داده شده است.

جدول ۲. مقدار بار عاملی استاندارد شده و آماره تی بین متغیرها

متغیر	مؤلفه‌ها	مقدار بار عاملی استاندارد شده	مقدار t	R2	آلفای کرونباخ	نتیجه
وضعیت موجود	نیروی انسانی	۰/۶۴۳	۱۱/۳۴۸	۰/۴۱۴	۰/۹۷۶	مطلوب
	سازمان	۰/۵۱۷	۸/۸۴۱	۰/۲۶۷		مطلوب
	زیرساختار	۰/۵۹۵	۱۰/۲۳۰	۰/۳۵۴		مطلوب
	مسائل اقتصادی	۰/۶۰۱	۱۰/۲۳۳	۰/۳۶۱		مطلوب
	فرهنگ	۰/۵۷۲	۹/۲۶۶	۰/۳۲۸		مطلوب
	مدیریت داده	۰/۵۷۸	۱۰/۰۵۴	۰/۳۳۴		مطلوب

همان‌طور که در جدول ۲ نمایان است، مقدار آماره تی وضعیت موجود با مؤلفه‌های خود معنادار و بیشتر از ۱/۹۶ برآورد شده است. ضریب آلفای کرونباخ وضعیت موجود (۰/۹۷۶) نیز بیشتر از ۰/۷ محاسبه شده است و نشان می‌دهد که پایایی سازه‌ها در سطح مطلوبی قرار دارد.

شاخص‌های برازش مدل

جدول ۳ شاخص‌های برازش مدل را نشان می‌دهد.

جدول ۳. شاخص‌های برازش مدل مرتبه دوم وضعیت موجود

x2/ df<3	RMSEA<0.08	CFI>0.9	IFI>0.9	GFI>0.8
۱/۱۲۹	۰/۰۱۸	۰/۹۸	۰/۹۸	۰/۸۳

مقدار شاخص کای اسکوئر (x2) بر درجه آزادی ۱/۱۲۹، شاخص نیکویی برازش (GFI) ۰/۸۳، شاخص برازش مقایسه‌ای (CFI) ۰/۹۸، شاخص برازش فزاینده (IFI) ۰/۹۸ و شاخص بسیار مهم ریشه دوم برآورد واریانس خطای تقریب یا همان ریشه میانگین مربعات خطای برآورد (RMSEA) ۰/۰۱۸ برآورد شده است. بنابراین شاخص‌های برازش مدل در حالت کلی بسیار مطلوب بوده و در سطح قابل‌قبولی قرار دارد. مدل عاملی تأییدی مرتبه دوم نشان می‌دهد که مؤلفه‌های ذیل به ترتیب بیشترین تا کمترین اهمیت را در وضعیت موجود دارند، که عبارت‌اند از:

۱) متغیر نیروی انسانی با میزان بار عاملی ۰/۶۴ و ضریب تعیین (R2) ۴۱ درصدی.

۲) متغیر مسائل اقتصادی با میزان بار عاملی ۰/۶۰ و ضریب تعیین (R2) ۳۶ درصدی.

۳) متغیر زیرساختار با میزان بار عاملی ۰/۵۹ و ضریب تعیین (R2) ۳۵ درصدی.
۴) متغیر مدیریت داده با میزان بار عاملی ۰/۵۸ و ضریب تعیین (R2) ۳۳ درصدی.
۵) متغیر فرهنگ با میزان بار عاملی ۰/۵۷ و ضریب تعیین (R2) ۳۲ درصدی.
۶) متغیر سازمان با میزان بار عاملی ۰/۵۲ و ضریب تعیین (R2) ۲۶ درصدی.

نتیجه

در این پژوهش وضعیت موجود مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش‌ابری در کتابخانه‌های دانشگاه‌های سطح اول کشور از دیدگاه کتابداران بررسی و مؤلفه‌های آن رتبه‌بندی شد. وضعیت موجود با میانگین ۲۳/۵۸۲ به‌طور معناداری در هر ۱۳ کتابخانه دانشگاهی سطح اول کشور در وضعیت نامطلوب قرار داشت. در رتبه‌بندی حاصل از تحلیل عاملی تأییدی به لحاظ قوی بودن مؤلفه نیروی انسانی با میزان بار عاملی ۰/۶۴ و ضریب تعیین (R^2) ۴۱ درصدی در رتبه اول قرار گرفت. نتایج مطالعه حاضر حاکی از آشنایی متوسط نیروی انسانی شاغل در آن کتابخانه‌ها با واژه‌های کلان‌داده و رایانش‌ابری بوده است. احمد، جیانگ مین و رافی[1] (۲۰۱۹)، در مقاله‌ای با عنوان «تحلیلی بر صلاحیت‌ها و مهارت‌های پیاده‌سازی تجزیه‌وتحلیل کلان‌داده‌ای در کتابخانه‌ها: یک مطالعه همبستگی» نشان دادند که همبستگی بسیار قوی بین صلاحیت‌ها و مهارت‌های موردنیاز کتابداران برای پیاده‌سازی تجزیه‌وتحلیل کلان‌داده‌ای وجود دارد. آنها ضمن بیان آگاهی بالای پاسخ‌دهندگان از مفهوم کلان‌داده، عنوان کردند که هم اکنون کتابداران درگیری‌های اولیه‌ای با فعالیت‌های مرتبط با کلان‌داده را دارند. آنها تأکید کردند که کلان‌داده برای پیاده‌سازی فعالیت‌های مختلف کتابخانه‌ای مانند انتخاب، مجموعه‌سازی حفاظت و غیره بسیار مفید است. قبادپور، نقشینه و ثابت‌پور (۱۳۹۱) در پژوهش خود پیش‌بینی کردند که در سال‌های آینده خدمات کتابداران بیشتر از امروز وابسته به کسب مهارت فناوری اطلاعات می‌باشد و چنان چه کتابداران خود را به کسب مهارت بالا در استفاده از فناوری آماده نسازند، دیری نمی‌پاید که مهندسان رایانه و یا فناوران اطلاعات جای خود را در کتابخانه‌ها باز می‌کنند و جایگاه کتابداران را به خطر می‌اندازند. در باب اهمیت مؤلفه نیروی انسانی و با هدف تحلیل ویژگی‌های فردی و حرفه‌ای پاسخگویان در این پژوهش با مؤلفه نیروی انسانی می‌توان این‌گونه اظهار داشت؛ همان‌طور که از نتایج حاصل از بررسی توزیع سن پاسخگویان نمایان است بیشترین تعداد پاسخگویان و در واقع کتابداران شاغل در کتابخانه‌های سطح اول در رده سنی ۲۹ تا ۳۹ سال

1.Ahmad, JianMing & Rafi

می‌باشند. اهمیت متغیر سن در این پژوهش به مسئله توان جسمی و ذهنی کتابداران در فراگیری فنون و روش‌ها و کسب مهارت‌های جدید در حوزه‌های کلان‌داده و رایانش ابری می‌باشد. جمعیت جوان شاغل در این دانشگاه‌ها خود نقطه قوت و نشان‌دهنده توانمندی کتابداران در آموزش و یادگیری و به‌روزرسانی خود با فناوری‌های روز دنیا و ارائه خدمات با بهره‌گیری از این آموزش‌ها است. متغیر دیگری که می‌تواند بر میزان آشنایی و استفاده از فناوری‌های کلان داده و ابر تأثیرگذار باشد، عامل سطح تحصیلات کتابداران شاغل در جامعه پژوهش است. در سال‌های اخیر به دلایلی مانند افزایش ظرفیت پذیرش در مقطع کارشناسی ارشد رشته علم اطلاعات و دانش‌شناسی، کتابداران دانشگاهی از لحاظ مدرک تحصیلی رشد چشمگیری داشته‌اند (اسفندیاری مقدم، حسن‌زاده و غیوری، ۱۳۹۰). علاوه بر افزایش ظرفیت در مقطع کارشناسی ارشد آن رشته می‌توان به دلایل دیگری مانند افزایش گرایش‌ها و حوزه‌های تخصصی در رشته، تلاش برای کسب رضایت تحصیلی و شغلی، رشد و افزایش نقش‌های اجتماعی و فرهنگی، درآمد، عدم رضایت از وضعیت موجود متقاضی و سایر موارد مثبت و منفی دیگر اشاره کرد. به هر روی افزایش سطح تحصیلات کتابداران در جامعه پژوهش حاضر، منجر به آشنایی بیشتر و توانمند شدن افراد با فناوری‌های کلان‌داده و رایانش ابری می‌گردد. می‌توان انتظار داشت که کتابخانه‌های سطح اول کشور با برخورداری از کتابدارانی به‌روزتر، توانمندتر، عالم‌تر قادر به تغییر نقش‌های سنتی خود و کسب مهارت‌های جدید هستند. کارمندان باسوادتر، خدمات بهتری را ارائه و پویایی و نقش کتابخانه‌ها را در عصر فناوری‌های نو بهبود بخشیده و به رونق حرفه کمک شایانی خواهند کرد.

در رتبه‌بندی حاصل از تحلیل عاملی تأییدی، مؤلفه مسائل اقتصادی با میزان بار عاملی ۰/۶۰ و ضریب تعیین (R^2) ۳۶ درصدی در رتبه دوم قرار گرفت. با عنایت به تطبیق متوسط بودجه کتابخانه‌های دانشگاه‌های سطح اول کشور با نیازهای فعلی و تأمین هزینه‌های لازم برای دیجیتال‌سازی منابع می‌توان مطابق با نتایج حاصل از پژوهش اذعان داشت که این دانشگاه‌ها قادر به تأمین هزینه‌های لازم برای زیرساختارهای موردنیاز فناوری‌های کلان‌داده‌ای مبتنی بر رایانش ابری و تأمین امنیت لازم برای آنها می‌باشند. اگرچه نتایج حاکی از هزینه‌های بسیار پایین برای آموزش کار با این فناوری‌ها برای کتابداران می‌باشد.

مؤلفه زیرساختار با میزان بار عاملی ۰/۵۹ و ضریب تعیین (R^2) ۳۵ درصدی در رتبه سوم جای گرفت. وجود زیر ساختارهای قانونی سخت‌افزاری و نرم‌افزاری مناسب در کتابخانه‌های دانشگاه‌های سطح اول راه را برای انجام اقدامات لازم برای بهره‌برداری از سرویس‌های ذخیره‌سازی، محاسباتی پردازشی، تحلیلی، بصری‌سازی از مجموعه سرویس‌های موجود برای کلان‌داده‌های مبتنی بر ابر هموار می‌کند. مؤلفه

مدیریت داده با میزان بار عاملی ۰/۵۸ و ضریب تعیین (R^2) ۳۳ درصدی در رتبه چهارم قرار گرفت. دسترسی به پایگاه‌های اطلاعاتی معتبر جهان در کتابخانه‌های دانشگاه‌های سطح اول کشور، کتابداران را با داده‌هایی که ویژگی‌های اصلی کلان داده یعنی حجم، سرعت و تنوع بالا را دارند، درگیر کرده است. لذا کتابداران در حال کار با چنین داده‌هایی هستند. این حجم عظیم نیازمند نگاهی یکپارچه و مدیریتی است که با استفاده از حاکمیت داده[1] که در اینجا حاکمیت کلان‌داده[2] مطرح می‌باشد، می‌توان به این مهم دست یافت. حاکمیت کلان‌داده، به مدیریت حجم انبوهی از داده‌های سازمان‌ها و بهره‌گیری از آنها در راستای فرایندهای تصمیم‌سازی با بهره‌گیری از ابزارهای تحلیلی گوناگون اشاره دارد. موسسه فناوری ماساچوست[3] پیرامون حاکمیت کلان‌داده چنین می‌گوید: «ایجاد یک مدل حاکمیت خوب، بستگی به سطح پختگی سازمان برای تصمیم‌گیری داده‌محور دارد». مسئله کلیدی حاکمیت کلان‌داده وابسته به برنامه‌های کاربردی می‌باشد که اطلاعات بسیاری را به‌صورت جریان پیوسته داده از منابع خارجی دریافت می‌کنند. بنابراین یک سیاست‌گذاری قابل‌قبول و شفاف در رابطه با نوع داده‌های ذخیره‌ای باید به‌وجود بیاید و همچنین باید مشخص شود که افراد مختلف با چه سرعتی به داده دسترسی پیدا کرده و سطح دسترسی داده به چه میزان باید باشد. نتایج این پژوهش، اقدامات بسیار اندکی را در زمینه حاکمیت داده را نشان داد. با توجه به این‌که در این پژوهش ذخیره‌سازی کلان‌داده‌ها مبتنی بر ابر[4] مدنظر بود، نتایج حاکی از عدم ذخیره‌سازی داده در سرویس‌های ابری در کتابخانه‌های دانشگاه‌های سطح اول بوده است. مؤلفه فرهنگ با میزان بار عاملی ۰/۵۷ و ضریب تعیین (R^2) ۳۲ درصدی در رتبه پنجم قرار گرفت. درک فرهنگ دسترسی آزاد به اطلاعات و به اشتراک‌گذاری آن توسط کتابداران از وظایف کارکردی کتابخانه‌ها محسوب می‌شود. نتایج در این زمینه حکایت از درک اکثریت کتابداران به دسترسی آزاد به داده و اطلاعات دارد. دسترسی آزاد به اطلاعات جزء حقوق شهروندی افراد در ایران محسوب می‌گردد و در اکثر کشورهای جهان نیز پذیرفته شده است (سلیمان دهکردی، افراسیابی، ۱۳۹۴). اغلب کتابداران اهمیت ارائه داده‌های تحلیل شده به مراجعین را درک کرده‌اند، اما درک مراجعین از دریافت چنین داده‌هایی در حد متوسط قرار داشت، که علت آن‌را می‌توان به عوامل متعددی از قبیل شناخت کم و عدم درک تأثیر استفاده از فناوری‌های جدید در امور پژوهشی، عدم درک ضرورت یک فرهنگ تصمیم‌گیری داده محور، عدم درک ضرورت به‌کارگیری فناوری‌های

1. Data Governance
2. Big Data Governance
3. Massachusetts Institute of Technology
4. Storage Based-Cloud

مدیریت کلان‌داده‌های مبتنی بر ابر برای رویارویی با حجم عظیم داده‌ها، عدم بهره‌مندی از خدمات کلان‌داده در کتابخانه‌ها و مواردی از این قبیل نسبت داد. از طرفی کتابداران با تجزیه‌وتحلیل کلان‌داده‌ها قادر به ارزیابی خدمات و برنامه‌های خود هستند. کیم و کوک[1] (۲۰۱۷) با بررسی یک برنامه کاربردی تجزیه‌وتحلیل کلان‌داده بر اساس روش‌های صورتک‌های چرنوف[2] به مقایسه و تطبیق خدمات و عملیات‌های انجام شده میان کتابخانه‌های عمومی لندن و سئول پرداختند. علاوه بر این، کتابداران در دانشگاه سنت فرانسیس لاورن[3] و کتابخانه دوروتی براون[4] از تجزیه‌وتحلیل گوگل[5] استفاده کردند تا دسترسی کاربران به منابع کتابخانه را از طریق دستگاه‌های تلفن همراه خود نشان دهند (Cottrell & Bell, 2016).

مؤلفه سازمان با میزان بار عاملی ۰/۵۲ و ضریب تعیین (R^2) ۲۶ درصدی در رتبه آخر قرار گرفت. وجود آگاهی و درک پایین مسئولین در ضرورت توسعه اهداف و راهبردها برای بهره‌گیری از فناوری‌ها مدیریت کلان‌داده مبتنی بر ابر در کتابخانه‌ها و عدم تمایل به تعامل و همکاری با سازمان‌های فعال در این حوزه، وجود محیط پایداری در کتابخانه‌های دانشگاه‌های سطح اول جهت کار با کلان‌داده‌ها و ارائه خدمات بر مبنای آنها را فراهم نمی‌کند. علی‌رغم اینکه برای دهه‌های متوالی، محققین نظریه‌های مختلفی را برای ناتوانی سازمان‌ها در همگام‌سازی خود با پیشرفت‌های کسب‌وکارهای جاری، سازگاری و مطابقت خود با نوآوری‌ها یا سازگاری با تغییرات محیطی ارائه داده‌اند (Raisch, Birkinshaw, Probst &Tushman, 2009; Del Giudice & Della Peruta, 2016; Marzi, Dabić, Daim, & Garces, 2017) و راه‌حل‌هایی از قبیل توجه به خلاقیت و نوآوری و حفظ کارمندان با قابلیت‌های بالا، ایجاد چابکی در سازمان‌ها، کسب رضایت مشتریان، رسیدگی به فرآیندها و عملیات‌ها، منابع مادی، مدیران و رهبران اثرگذار و سایر موارد مطرح شده در این زمینه را ارائه کرده‌اند؛ مطابق نتایج این پژوهش شاهد عدم توجه مسئولین به خلاقیت و نوآوری در کتابخانه‌ها، جذب افراد غیرمتخصص، عدم همکاری و تعامل با سازمان‌ها و شرکت‌های متخصص در آن حوزه مربوطه و نبود قوانین و مقررات شفاف جهت دسترسی و به اشتراک‌گذاری داده‌ها و رعایت حریم خصوصی، مالکیت فکری هستیم. سازمان‌ها، سیاست‌ها و تصمیمات آنها، عامل‌هایی حیاتی برای شروع به‌کارگیری و استفاده از یک فناوری هستند. سازمان‌ها و اقدامات آنها، تمامی مباحث مرتبط با نیروی انسانی، هزینه‌ها، زیر ساختارها و فرهنگ‌سازی را در به‌کارگیری فناوری‌های نوظهور به‌شدت تحت

1. Kim & Cooke
2. Chernoff Face
3. St. Francis' LaVerne
4. Dorothy Brown Library
5. Google Analytics

تأثیر قرار می‌دهند. همچنین سیاست‌ها، به‌عنوان حوزه روابط قدرت نیز در نظر گرفته می‌شوند و با سازمان‌ها در تعامل منسجم با یکدیگر هستند. به همین دلیل در یک تحلیل ترکیبی مدنظر قرار می‌گیرند. بنابراین نقش حیاتی ساختارهای سازمانی و عملکرد آنها در تعیین موفقیت و یا شکست فناوری مشخص می‌باشد. با توجه به ضرورت به‌کارگیری الزامات قانونی مطرح برای کلان‌داده‌ها در موارد رعایت حریم شخصی، مالکیت معنوی داده‌ها و حفاظت از داده‌ها و با عنایت به این امر که تاکنون در سراسر دنیا یک قانون یا آئین‌نامه منسجم مختص کلان‌داده‌ها ارائه نشده است و بحث کلان‌داده‌ها و الزامات حقوقی آن همچنان بحث باز کشورها می‌باشد (معین، ۱۳۹۵). بنابراین می‌توان روال‌ها و قوانین لازم برای به‌کارگیری کلان‌داده‌ها در کتابخانه را نیز در آن گنجاند.

پیشنهادهایی برای پژوهشگران آینده

قطعاً انجام یک پژوهش قادر به شناسایی تمامی ابعاد برای شناخت کامل وضعیت موجود مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌ها نمی‌باشد و انجام پژوهش‌های ذیل می‌تواند در جهت تکمیل و رفع عیوب پژوهش حاضر مؤثر باشد. بنابراین انجام پژوهش‌های زیر پیشنهاد می‌شود.

۱) از آنجایی که پایه و اساس کلان‌داده‌ها کار با داده‌های دیجیتالی است، پیشنهاد می‌شود پژوهشی درباره مدیریت کلان‌داده‌ها در بستر ابر در کتابخانه‌های دیجیتال ایران انجام شود.

۲) در پژوهشی به تبیین نقش‌ها و مهارت‌های جدید و ضروری کتابداران در عصر کلان‌داده پرداخته شود.

۳) با عنایت به ضرورت درک و توجه ویژه در سطوح کلان مدیریتی، پژوهش‌های زمینه‌ای درباره میزان شناخت و تخصص مسئولین و مدیران با فناوری‌های کلان‌داده و رایانش ابری انجام گیرد.

۴) در پژوهشی مستقل به شناسایی چارچوب معماری کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در کتابخانه‌های دانشگاه‌های سطح اول کشور پرداخته شود.

۵) در پژوهشی مستقل به شناسایی زیرساخت‌های سخت‌افزاری و نرم‌افزاری مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در ایران پرداخته شود.

۶) در پژوهشی مستقل به شناسایی و معرفی برنامه‌های کاربردی مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری در سایر کشورهای جهان پرداخته شود.

۷) در پژوهشی مستقل به شناسایی فنون تحلیل کلان‌داده‌های کتابخانه‌ای و مدیریت آنها مبتنی بر رایانش ابری پرداخته شود.

۸) در پژوهشی مستقل به شناسایی زیرساخت امنیتی و مباحث مربوط به قانون حق مؤلف در مدیریت کلان‌داده‌های کتابخانه‌ای مبتنی بر رایانش ابری پرداخته شود.

منابع

آقایی میرک آباد، اعظم، آهنگر، عادله، آقایی میرک آباد، الهه و حسن‌پور، طلعت (۱۳۹۵). *سیستم‌های مدیریت کتابخانه کلان داده‌های مبتنی بر مدل ابری*. به کوشش میترا قیاسی، سید علی اصغررضوی، صفیه طهماسبی لیمونی، مجموعه مقالات همایش ملی مدیریت داده‌های عظیم در علم اطلاعات و دانش‌شناسی (ص ۶۷–۷۸). بابل: دانشگاه آزاد اسلامی.

ارل، توماس، ختاک، وجید و بوهلر، پاول. (۱۳۹۵). *اصول داده‌های بزرگ مفاهیم، پیشران‌ها و تکنیک‌ها*؛ ترجمه رشیدی و دیگران. تهران: انتشارات آتی نگر.

اسفندیاری مقدم، علیرضا، حسن‌زاده، محمد و غیوری، زینب (۱۳۹۰). تحلیل عامل‌های اثرگذار بر تولیدات علمی زنان ایرانی در آی. اس. آی. *پژوهشنامه کتابداری و اطلاع‌رسانی*،۱ (۱)،۱۰۹–۱۳۴. https://dx.doi.org/10.22067/riis.v1i1.8487

چن، مین، ماو، شیون، ژانگ، ین و ام.لیونگ، ویکتور سی (۱۳۹۴). *کلان داده‌ها: فناوری‌ها*؛ ترجمه امیر مسعود رحمانی، رقیه توکلیان. بابل: انتشارات علوم رایانه.

حبیبی، آرش (۱۳۹۶). *مدل یابی معادلات ساختاری و تحلیل عاملی (آموزش کاربردی نرم‌افزار LISREL)*. تهران: جهاد دانشگاهی.

درمنده، مژده، نوروزی، علیرضا و اسمعیلی گیوی، محمدرضا (۱۳۹۷). فرصت‌های مدیریت داده‌های بزرگ در کتابخانه‌ها و مراکز اطلاع‌رسانی: واکاوی ساختاری–تفسیری و ارائه راهکار. *پژوهشنامه پردازش و مدیریت اطلاعات*. ۱۳۹۷; ۳۴ (۲): ۸۴۱–۸۷ http://jipm.irandoc.ac.ir/article-1-3866-fa.html

ساسینسکی، بری ا.(۱۳۹۰). *مرجع کامل رایانش ابری*. ترجمه نوید فرخی. تهران: علوم رایانه.

سلیمان دهکردی، الهام؛ افراسیابی، علی (۱۳۹۴). حق دسترسی آزاد به اطلاعات در نظام حقوقی بین‌المللی و ملی. *مطالعات بین‌المللی پلیس*، ۶(۲۱)، ۷۵–۱۰۰. http://interpol.jrl.police.ir/article_12886.html

شهرزادی، لیلا (۱۳۹۱). *کارگاه آشنایی با منابع اطلاعاتی علمی الکترونیکی*. دانشگاه علوم پزشکی اصفهان. بازیابی شده از: http://lib.mui.ac.ir/sites/lib.mui.ac.ir/files/amozesh/ashen.pdf

صراف‌زاده، مریم (۱۳۹۴). مدیریت کلان‌داده‌های پژوهشی: نقشی نوین برای کتابخانه‌های دانشگاهی. *فصلنامه نقد کتاب اطلاع‌رسانی و ارتباطات*. ۶، ۲۶۵ – ۲۷۴. http://icbr.faslnameh.org/article-1-183-fa.html

قبادپور، وفا، نقشینه، نادر و ثابت پور، افسون (۱۳۹۱). از رایانش ابری تا کتابخانه ابری و ارائه پیشنهاد طراحی کتابخانه با الگوی رایانش ابری. *پردازش و مدیریت اطلاعات (علوم و فناوری اطلاعات)*، ۲۸ (۴)،۸۵۹–۸۷۷. http://jipm.irandoc.ac.ir/article-1-2384-fa.html

معین، محمد شهرام (۱۳۹۵). *الزامات قانونی کلان داده‌ها. پروژه تدوین نقشه راه کلان داده‌ها*. تهیه‌کننده فاطمه کسائی. تهران: مرکز تحقیقات مخابرات ایران.

Adebayo, O. A., Ahmed, Y. O., & Adeniran, R. T. (2018). The role of ict in provision of library services a panacea for sustainable development in Nigeria. *Library Philosophy & Practice*. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=5305&context=libphilprac (accessed 11 July 2019).

Ahmad, K., JianMing, Z., & Rafi, M. (2019). An analysis of academic librarians' competencies and skills for implementation of Big Data analytics in libraries. *Data Technologies and Applications*. https://doi.org/10.1108/DTA-09-2018-0085

Al-Daihani, S. M., & Abrahams, A. (2016). A text mining analysis of academic libraries' tweets. *The journal of academic librarianship*, 42(2), 135-143. https://doi.org/10.1016/j.acalib.2015.12.014

Avram, M. G. (2014). Advantages and challenges of adopting cloud computing from an enterprise perspective. *Procedia Technology*, 12(0), 529-534. https://doi.org/10.1016/j.protcy.2013.12.525

Ball, R. (2019). Big data and their impact on libraries. *American Journal of Information Science and Technology*, 3(1), 1-9. http://dx.doi.org/10.11648/j.ajist.20190301.11

Cottrell, T., & Bell, B. (2016). When to say when: Using big data to support mobile communications. *College & Undergraduate Libraries*, 23(3), 315-322. https://doi.org/10.1080/10691316.2015.1023878

Del Giudice, M., & Della Peruta, M. R. (2016). The impact of IT-based knowledge management systems on internal venturing and innovation: a structural equation modeling approach to corporate performance. *Journal of Knowledge Management*. 20(3). https://doi.org/10.1108/JKM-07-2015-0257

Fister, B. (2015). Big data or big brother? data, ethics, and academic libraries. *Library Issues: Briefings for Faculty and Administrators*, 35(4), 293-314.
Available at: http://barbarafister.com/LIbigdata.pdf (accessed 15 December 2019).

Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. *IDC rview*, 1142(2011), 1-12. http://www.kushima.org/wp-content/uploads/2013/05/DigitalUniverse2011.pdf (accessed 25 June 2018).

Gordon-Murnane, L. (2012). Big Data: A big opportunity for librarians. *Online (Weston, CT)*, 36(5), 30-34. http://eprints.rclis.org/34349/1/final-proceeding-iicclist16-2s-view-50-51.pdf (accessed 21 Apr 2019).

Hamad, F., Fakhuri, H., & Abdel Jabbar, S. (2020). Big Data Opportunities and Challenges for Analytics Strategies in Jordanian *Academic Libraries. New Review of Academic Librarianship*, (just-accepted), 1-19. https://doi.org/10.1080/13614533.2020.1764071

Head, S. R., Komori, H. K., LaMere, S. A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D. R., & Ordoukhanian, P. (2014). Library construction for next-generation sequencing: overviews and challenges. *Biotechniques*, 56(2), 61-77. https://www.future-science.com/doi/full/10.2144/000114133 (accessed 3 Apr 2019).

Heidorn, P. B. (2011). The emerging role of libraries in data curation and e-science. *Journal of*

Library Administration, 51(7-8), 662-672. https://doi.org/10.1080/01930826.2011.601269

Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013, January). Big data: Issues and challenges moving forward. InSystem Sciences (HICSS), 2013 46th Hawaii International Conference on (pp. 995-1004). *IEEE*. https://doi.org/10.1109/HICSS.2013.645

Keil, D. E. (2014). Research data needs from academic libraries: The perspective of a faculty researcher. *Journal of Library Administration*, 54(3), 233-240. https://doi.org/10.1080/01930826.2014.915168

Kim, S., & Choi, M. S. (2016, February). Study on data center and data librarian role for reuse of research data. In 2016 8th International Conference on Knowledge and Smart Technology (KST) (pp. 303-308). *IEEE*. https://doi.org/10.1109/KST.2016.7440517

Kim, Y. S., & Cooke, L. (2017). Big data analysis of public library operations and services by using the Chernoff face method. *Journal of Documentation*. https://doi.org/10.1108/JD-08-2016-0098

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META group research note*, 6(70), 1. http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf (accessed 8 may 2019).

Li, J., Lu, M., Dou, G., & Wang, S. (2017). Big data application framework and its feasibility analysis in library. *Information Discovery and Delivery*. https://doi.org/10.1108/IDD-03-2017-0024

Lin, C. H., Lin, I. C., Roan, J. S., & Yeh, J. S. (2012). Critical factors influencing hospitals' adoption of HL7 version 2 standards: An empirical investigation. *Journal of medical systems*, 36(3), 1183-1192. https://doi.org/10.1007/s10916-010-9580-2

Lyon, L., Wright, S., Corti, L., Edmunds, S., & Bennett, F. (2013). *What is a data scientist?* Panel presented at 2013 International Digital Curation Conference, January 14-17, 2013, Amsterdam, Netherlands.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Byers, A.H., 2011. *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

Marshall, J. G., Sollenberger, J., Easterby-Gannett, S., Morgan, L. K., Klem, M. L., Cavanaugh, S. K., … & Hunter, S. (2013). The value of library and information services in patient care: results of a multisite study. *Journal of the Medical Library Association: JMLA*, 101(1), 38. https://dx.doi.org/10.3163%2F1536-5050.101.1.007

Marzi, G., Dabić, M., Daim, T., & Garces, E. (2017). Product and process innovation in manufacturing firms: a 30-year bibliometric analysis. *Scientometrics*, 113(2), 673-704. https://doi.org/10.1007/s11192-017-2500-1

Pryor, G., & Donnelly, M. (2009). Skilling up to do data: whose role, whose responsibility, whose career. *International Journal of Digital Curation*, 4(2), 158-170. https://doi.org/10.2218/ijdc.v4i2.105

Raisch, S., Birkinshaw, J., Probst, G., & Tushman, M. L. (2009). Organizational ambidexterity: Balancing exploitation and exploration for sustained performance. *Organization science*, 20(4), 685-695. https://doi.org/10.1287/orsc.1090.0428

Rani, B. R. (2016). *Big Data and Academic Libraries*. In International conference on Big Data and knowledge discovery. Indian Statistical Institute.

Qutab, S., Bhatti, R., & Ullah, F. S. (2014). Adoption of ICT's for library operations and

services: A comparison of public and private university libraries of Pakistan. *Library Philosophy and Practice*, 0_1. https://digitalcommons.unl.edu/libphilprac/1106 (accessed 21 June 2017).

Reinhalter, L., & Wittmann, R. J. (2014). The Library: Big Data's Boomtown: Edited by Rick J. Block. *The Serials Librarian*, 67(4), 363-372. https://doi.org/10.1080/0361526X.2014.915605

Swan, A., & Brown, S. (2008). *The skills, role and career structure of data scientists and curators: An assessment of current practice and future needs*. Truro: Key Perspectives Ltd. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.147.8960&rep=rep1&type=pdf

Teets, M., & Goldner, M. (2013). Libraries' role in curating and exposing big data. *Future internet*, 5(3), 429-438. https://doi.org/10.3390/fi5030429

Witt, M. (2012). Co-designing, co-developing, and co-implementing an institutional data repository service. *Journal of Library Administration*, 52(2), 172-188. https://doi.org/10.1080/01930826.2012.655607

Yuhong, X. I. E. (2018). Utilization and Reconstruction of Digital Resources in University Libraries in the Perspective of Big Data Based on the Statistics of Digital Resources Utilization in the Changbei University Alliance Library. *Modern Information Technology*, (6), 61. https://en.cnki.com.cn/Article_en/CJFDTOTAL-XDXK201806061.htm

Big Data Management based on Cloud Computing in the Libraries of First Level Universities in Iran

Seyedeh Elham Hashemi Bonjar*
Soraya ziaei,
Maryam Salami

Mashhad Payam Noor University

Abstract

Introduction: Big data is a set of data that, with its special features, cannot be stored, managed and processed by conventional software systems. Cloud computing is a collection of virtualized resources with easy access and usability. Cloud-based storage technology is able to effectively manage big data. The purpose of this study is to determine the current status of libraries̓ big data management based on cloud computing and its Component ranking in the libraries of first level universities in Iran

Methodology: The present study is an applied and in terms of nature, method and control of variables is a descriptive research. The statistical population includes all librarians working in the libraries of first level universities in Iran, 520 people. The

* . elnaz_hashemi2933@yahoo.com

questionnaire was provided to the respondents electronically and a total of 393 questionnaires were received and analyzed. Data collection tool is a researcher-made questionnaire that examines previous studies of six main components affecting the libraries̓ big data management based on cloud computing in the libraries of first level universities in Iran that including manpower, organization, infrastructure, economic issues, culture and data management was identified, which was confirmed through confirmatory factor analysis. To analyze data from descriptive and inferential statistics with software Spss 22 and Amos 22 were used.

Findings: The current situation of library data management based on cloud computing in the libraries of Iran's top-level universities was considered unfavorable from the librarians' point of view, with an average of 23.582 and a standard deviation of 7.655 equal to -0.41730. Ranking the components of the current situation from strength to weakness, respectively 1. Manpower With a factor loading rate of 0.64, 2. Economic issues with a factor loading rate of0.60, 3. Infrastructure With a factor loading rate of0.59, 4. Data Management With a factor loading rate of0.58, 5. Culture With a factor loading rate of 0.57 6. Organizations With a factor loading rate of 0.52 are located.

Conclusion: Librarians are practically dealing with big data because of their day-to-day dealings with very large volumes of data. Matching the average budget of these libraries with the current needs and providing the necessary funds for the digitization of resources and the existence of appropriate legal hardware and software infrastructure, it can pave the way for the necessary steps to take advantage of cloud-based storage services. Most librarians, while understanding the culture of free access to information, understand the importance of providing analyzed data to clients, but clients' perception of receiving such data is moderate. Low awareness and understanding of officials and unwillingness to interact and cooperate with organizations active in this field, does not provide a stable environment to work with big data and provide services based on them.

Keywords: Big Data Management, Cloud Computing, Academic Libraries, Librarians